Destek Vektör Makineleri ile Yaramaz Elektronik Postaların Filtrelenmesi

Spam e-mail Filtering Using Support Vector Machine

E. U. Küçüksille [1] ve N. Ateş [2]

[1]Süleyman Demirel Üniversitesi, Isparta/Turkey, ecirkucuksille@sdu.edu.tr
[2]Süleyman Demirel Üniversitesi, Isparta/Turkey, nurullahates@sdu.edu.tr

Özetçe

Elektronik postanın (e-posta) hızlı ve kolay bir haberleşme aracı olması, insanlar tarafından iletişimde yoğun şekilde kullanılmasına neden olmaktadır. E-postanın bu özellikleri; reklam yapmak, fikirlerini insanlara duyurmak ve çeşitli istismar çalışmaları yapmak isteyen insanlar için bir cazibe merkezi olmasını sağlamaktadır. Bu tür girişimler de bir güvenlik zafiyeti oluşturmaktadır. Temelleri Vladimir N. Vapnik tarafından atılan bir makine öğrenme algoritması olan destek vektör makineleri 1995'den itibaren sınıflandırma ve eğri uydurma problemlerinde başarılı sonuçlar vermiştir. Bu çalışmada, destek vektör makineleri kullanılarak yaramaz (istenmeyen) e-postaların filtrelenmesi işlemi gerçekleştirilmiştir.

Anahtar Sözcükler: *Yaramaz (İstenmeyen e-posta) destek vektör makinesi, filtreler, öznitelik, çekirdek fonksiyonları.*

Abstract

Electronic mail (e-mail) is a communication medium that is fast and easy to use, making people to use it frequently. These features of e-mail cause it to be the center of attraction for the people who want to advertise, share their ideas with others or involve in malicious activities. Thus, these kind of activities create security threats. Support vector machines are machine learning algorithms that are developed by Vladimir N. Vapnik and produce successful results in several application domains such as clustering and curve fitting problem. In this study, we carry out unwanted (spam) e-mail filtering using support vector machine.

Keywords : *spam, support vector machine, filters, feature, kernel functions.*

1. GİRİŞ

Haberleşmenin temeli e-posta kullanımına dayanmaktadır. Bu popüler aracı kullanan insan sayısı çok fazla olduğundan birçok insan, grup veya şirketler, insanlara seslerini bu araç vasıtasıyla duyurmaya çalışmaktadır. Bir çalışmaya göre bir şirket ağına gelen mesajların %10'nunu yaramaz (istenmeyen) postalar (YP) oluşturmaktadır[12]. İnternet vasıtasıyla mesaj alma talebinde bulunmadığı halde çok sayıda kişiye gönderilen bu mesaj veya mesajlara yaramaz e-posta (YP) denir. YP'nin kullanılmasının en önemli avantajı, kısa bir süre içerisinde çok sayıda insana ulaşılmasıdır. YP'nin en belirgin zararları ise internet bant genişliğini doldurması ve zaman israfıdır. YP genellikle ticari amaçlı olmasının yanında siyasi bir propaganda ya da kamuoyu araştırması yapmak amacıyla gönderilmiş e-postalar da olabilmektedir. Bir e-posta adresine YP gelmesi için o adresin YP göndericinin eline geçmesi gerekmektedir. YP göndericileri adresleri ele geçirmek için ağı tarayan botlar kullanmaktadırlar. Arama motorlarının kullandığı türden olan bu botlar tüm siteleri veya formları tarayarak korumasız şekilde tutulan tüm bilgileri almaktadır. Kullanıcılar, üye oldukları (kişisel bilgilerini ve e-posta adreslerini verdikleri) her site ve forumlar üzerinden bilmeden YP sektörünü beslemektedirler.

Destek vektör makineleri(DVM) makine öğrenmesi gerektiren bir çok alanda başarılı sonuçlar vermiştir. YP'lerin filtrelenmesi konusunda da bir çok çalışmada farklı yöntemler ile birleştirilerek kullanılmıştır. Rafiqul ve Zhou çalışmalarında DVM

tabanına bağlı yenilikçi bir YP filtreleme çalışması yapmışlardır [7]. Zhiyang ve arkadaşları çalışmalarında Dvm'ye dayanarak ağ YP'lerinin araştırılması üzerine bir çalışma yapmışlarıdır [11]. Androutsopoulos ve arkadaşları çalışmalarında ağırlıklı Dvm kullanılarak YP filtreleme yöntemi kullanmışlardır [4]. Bu ve benzeri çalışmaların sonuçları DVM'nin YP filtreleme üzerine başarılı olduğunu göstermiştir. Bu çalışmada da öznitelik kümesi daha dikkatli seçildiğinde Dvm'nin performansının dikkate değer bir ölçüde arttığı görülmüştür.

2. ÖZNİTELİK ÇIKARTMA VE SEÇME

Öznitelik çıkartma, bazı kriterlere dayanarak mevcut bilgilere bir dönüştürme işlemi uygulayarak yeni bir öznitelik uzayı oluşturmak iken *öznitelik seçme*, mevcut öznitelikler arasından, bazı kriterlere dayanarak öznitelik seçme yani o örneği temsil edebilecek en iyi özniteliği seçmektir.

Her e-posta için en önemli parça onun içeriğidir (kelimeleridir). Bu yüzden YP filtreleme bir metin sınıflandırma problemidir. Metin sınıflandırma problemlerinin özniteliği de kelimelerdir. Öznitelik seçerken dikkat edilmesi gereken durumlar vardır. Örneğin, bir kelime hem YP hem de normal e-posta(np) da çok geçiyor veya ikisinde de az geçiyor ise o iyi bir öznitelik değildir. Kullanılan dil de öznitelik seçimini etkimektedir. Örneğin, İngilizce bir kelime olan 'got', 'get' ile benzer anlam ifade eder. Diğer bir örnek Türkçe bir kelime olan 'aldı' 'alır' ile benzer anlamlara sahYPtir. Bu iki örnekte de görüldüğü gibi diller arasında bir farklılık mevcuttur. Öznitelik seçerken bu farklılıklara dikkat etmek gerekmektedir. En az öznitelik sayısı ile YP tespit edilmesi hesaplama zorluğu ve zaman açısından önemlidir. Eğer öznitelik seçimi uygun bir şekilde yapılmaz ise öznitelik vektörünün boyutu artacak ve bu da hesaplama maliyetini arttıracaktır.

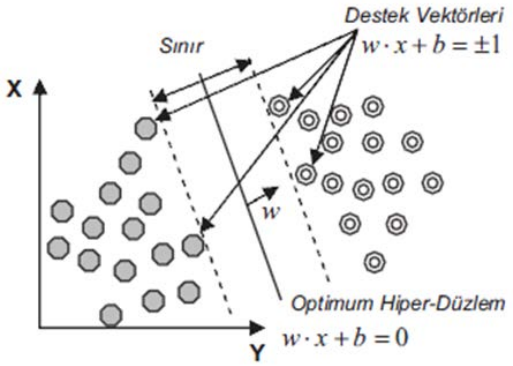
Şekil 3'te girdi uzayından özellik uzayına bir dönüşüm gösterilmektedir. Bu olay bir öznitelik çıkartma işlemidir.

3. DESTEK VEKTÖR MAKİNELERİ (DVM)

"Klasik istatistik, doğru modelin formunun bilindiğini varsayıp, amacı modelin parametrelerini belirlemek olarak görürken; istatistiksel öğrenme teorisi modelin formunun bilinmediğini kabul etmekte ve doğru olabilecek modeller arasından en iyi modelin bulunmasını hedeflemektedir"[9]. Destek vektöre makineleri (DVM), istatistiksel öğrenme teorisi ve yapısal riski en aza indirme ilkesine dayanan, sınıflandırma ve eğri uydurma problemlerinin çözümü amacıyla Vapnik tarafından ortaya atılmış bir öğrenme yöntemidir [10]. Bu öğrenme yöntemi denetimli öğrenme yöntemi kapsamına girer. Denetimli öğrenmede, eğitim aşamasında verilerin sınıf etiketleri yani hangi sınıfa ait oldukları bellidir. DVM' nin asıl yaptığı iş kendine girdi olarak gelen verileri 2 sınıfa ayırmaktır. Bu çalışmada DVM; doğrusal ayrılma, tam olarak doğrusal ayrılamama ve doğrusal ayrılamama olmak üzere 3 ana başlıkta incelenmiştir.

a. Doğrusal ayrılma

N adet elemandan oluşan eğitim veri kümesi $D = \{(\vec{x}_i, y_i), i = 1...N\}$ olduğu kabul edilirse buradaki $y_i \in \{-1,+1\}$ sınıf etiketi, $\vec{x}_i \in R^n$ olup n boyutlu uzayda herhangi bir örnektir. $f(\vec{x}) = \vec{w}^T \vec{x} + b$ ifadesinde ki $\vec{w}^T$ karar fonksiyonun normalini, $\vec{x}$ ifadesi bu doğru üzerinde bulunan noktaları, b ise eğilim değerini göstermektedir. Amaç $\vec{w}^T$ ve b' yi eğitim verileri yardımıyla bulmaktır, yani sistemi eğitmektir. Tüm destek vektör makinelerinde amaç Şekil 1'de olduğu gibi verileri 2 sınıfa ayırmaktır. Şekil 1'de kesikli çizgiler ile ifade edilen doğrular üzerindeki vektörler destek vektör olarak isimlendirilir ve yumuşak ayrım çizgisi bu vektörler üzerinden geçer. İki yumuşak ayrım çizgisinin ortasındaki doğru ise sert ayrımıdır ve $f(\vec{x}) = \vec{w}^T \vec{x} + b = 0$ fonksiyonuyla çizilir.



Şekil 1: Doğrusal olarak ayırma

$f(\vec{x}) = \vec{w}^T \vec{x} + b$ ifadesindeki $\vec{w}^T$ ve $\vec{x}$ vektörel büyüklük olup, (1)' in sade halini ifade eder.

$$f(\vec{x}) = w_1 x_1 + w_2 x_2 + w_3 x_3 + ... + w_n x_n + b \quad (1)$$

$f(\vec{x}) = \vec{w}^T \vec{x} + b \geq 1$ durumunda $y_i = 1$ ve $f(\vec{x}) = \vec{w}^T \vec{x} + b \leq -1$ durumunda $y_i = -1$ dir. İki fonksiyonu $y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1$ şeklinde kısalta biliriz. Bu ayırma işleminin püf noktası *sınır* değerini maksimum yapmak ve bu sayede en iyi ayrıma sahYP olmaktır. Veri setini sınıflara ayırabilecek sonsuz sayıda çoklu düzlem çizilebilmesine karşın, amaç bilinmeyen veri seti ile karşılaşıldığında sınıflama hatasını en küçük yapacak aşırı düzlemi seçmektir. Bunun için maksimum sınırlı aşırı düzlem tekniği önerilmiştir. *Sınır* değerinin büyüklüğü genelleme kabiliyetini arttırır. X_1 değeri $f(\vec{x}) = \vec{w}^T \vec{x} + b = 1$ fonksiyonu üzerinde bir nokta ve x_3 değeri $f(\vec{x}) = \vec{w}^T \vec{x} + b = -1$ fonksiyonu üzerinde bir noktadır. Sınır değerini bulmak için,

$$\vec{w}^T x_1 + b = +1 \quad (2)$$

$$\vec{w}^T x_3 + b = -1 \quad (3)$$

(3)' ü -1 ile çarpılıp (2) ile toplanYP $x_1 = x_3 + \lambda.w$ ifadesi eşitlikte yerine yazılırsa, bunun sonucunda $\lambda = 2/\vec{w}^2$ ifadesi bulunur. Hedef, λ değerini maksimum yapmak olduğu için $1/\lambda$ ifadesi minimum olmalıdır. Buna bağlı sınırlama ise $y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1 \geq 0$, $y_i \in \{-1, +1\}$ dir. Optimizasyon problemi, verilen bazı kısıtlamalar altında bir fonksiyonun maksimumunu ya da minimumunu bulmadır [3]. Bu optimizasyon problemi Lagrange denklemleri kullanılarak çözülebilir ve sonuçta,

$$L(\mathbf{w}, b, \mathfrak{9}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N} \alpha_i [y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1]$$
$$, \alpha_i \geq 0, \forall i \quad (4)$$

denklemi oluşur. Bu problem "Karush–Kuhn–Tucker (KKT)" in (5) ve (6) daki şartları kullanılarak çözülür.

$$\frac{\partial L}{\partial w_j} = 0, \quad \forall j \quad (5)$$

$$\frac{\partial L}{\partial w_j} = 0, \quad \forall j \quad (6)$$

(5) ve (6) eşitlikleri kullanılıp problem çözülünce (7)' deki formül oluşur ve bu formül için kısıtlar

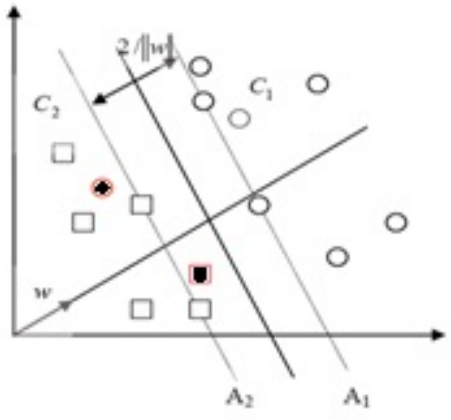
$\sum_{i=1}^{N} \alpha_i y_i = 0, \quad \alpha_i \geq 0, \quad \forall i$ dir.

$$L(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (7)$$

$L(\alpha)$ optimizasyon problemi standart 2. dereceden programlama teknikleriyle çözülür.

b. Tam olarak doğrusal ayrılamama

Veriler bazı durumlarda % 100 performansla ayrılamayabilir. Verilerin sınır içerisine düştüğü (Şekil 2'deki içi dolu kare) durum ayrılamama ve sert ayrılma çizgisinin karşı tarafına düştüğü (Şekil'2 deki içi dolu daire) durum da yanlış ayrılma olarak adlandırılır. Bu durumlarda, doğruların minimum hata ile ayırma sağlayacak şekilde ayarlanması gerekir.



Şekil 2: Tam olarak doğrusal ayıramama

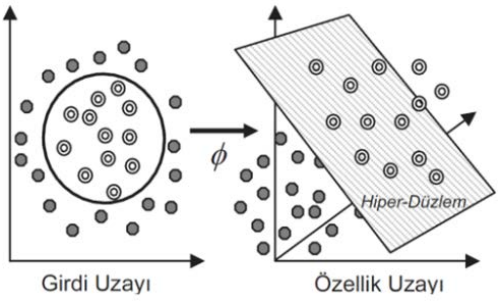
Yukarıda anlatılan sorunların çözümü için $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$ kısıt değeri (8)'deki gibi olur.

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i \tag{8}$$

Buradaki ξ_i gevşek değişken olarak isimlendirilir. ξ_i gevşek değişkeni bir $\mathbf{x}_i$ örneğinin sınırdan olan sapma uzaklığıdır. Ayrılamazlık durumunda ξ_i , $0 \leq \xi_i < 1$ ve yanlış sınıflandırma durumunda ξ_i , $\xi_i \geq 1$ olur. Doğrusal ayrılamama durumunda aynı doğrusal ayrılmada olduğu gibi maksimum *sınır* bulunmaya çalışılır. Sapma değerlerinin eklenmesiyle minimum formülasyon $\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i \xi_i^2$ şeklini alır. Buradaki 'C' ($0 < C < \infty$) düzenleme parametresi olarak isimlendirilir ve bu değer sınır genişliğini doğrudan etkiler. Tam olarak doğrusal ayrılamama optimizasyon problemi doğrusal ayrılabilen durumda olduğu gibi Lagrange denklemi ve KKT şartlarıyla çözülür. Denklemin çözümleri yapılınca tam olarak doğrusal ayrılama yani yumuşak ayrılma $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i$ denkleminin verileri elde edilmiş olunur ve artık sisteme gelen veriler bu formüle göre sınıflandırabilir.

c. Doğrusal ayrılamama (Çekirdek fonksiyonlar)

Gerçek dünyada veriler çoğunlukla doğrusal olarak ayrılamazlar, böyle durumlarda destek vektör makinesi Şekil-3 deki gibi girdi uzayını daha yüksek boyutlu bir uzaya taşıyarak burada doğrusal ayırma işlemelerini uygular.



Şekil 3: Doğrusal ayıramama

Girdi uzayının daha yüksek boyutlu uzaya taşınması için $\Phi(\mathbf{x})^T\Phi(\mathbf{x}_i) = K(\vec{x}, \vec{x}_i)$ şeklinde bir kernel fonksiyonu kullanılır. Bu bilgiden yola çıkarak lagrange fonksiyonu (9) daki gibi ifade edilir.

$$\sum_i \alpha_i - \frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j y_i y_j \Phi(\mathbf{x})^T\Phi(\mathbf{y}) =$$

$$\sum_i \alpha_i - \frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j y_i y_j K(\mathbf{x}, \mathbf{y})$$

(9)

Çekirdek düzenlemesi yapılarak dönüştürülmüş uzaydaki Φ (x) vektörü yerine girdi uzayındaki verilerden oluşan bir çekirdek fonksiyon oluşturularak işlemler gerçekleştirilir. Çekirdek fonksiyon kullanarak iç çarpım hesaplamak, dönüştürülmüş nitelik seti Φ (x) kullanarak hesaplamaya kıyasla daha kolaydır ve maliyeti düşüktür [2].

Doğrusal olmayan DVM'de kullanılan çekirdek fonksiyonları *Mercer Teoremi* olarak bilinen matematiksel bir kurala uymak zorundadırlar. Bu kural yüksek boyutta çalışılırken çekirdek fonksiyonların her zaman iki girdi vektörünün iç çarpımı şeklinde ifade edilmesini sağlamaktadır[8]. En çok kullanılan 2 çekirdek fonksiyonu *gauss* (10)' da ve *polynomial* (11)'de gösterilmiştir.

$$K(\vec{x}, \vec{x}_i) = e^{-\gamma \|(x - x_i)\|^2} \quad (10)$$

$$K(\vec{x}, \vec{x}_i) = ((x, x_i) + 1)^d \quad (11)$$

2 boyuttan 6 boyuta geçiş yapan polynomial çekirdek fonksiyonunun açılımı (12) de ve onun iç çarpan fonksiyonu (13) de görülmektedir.

$$K(\vec{x}, \vec{x}_i) = (1 + x^T . x_i)^2 = 1 + {x_1}^2 . {x_{i1}}^2 + 2 . x_1 . x_{i1} \; x_2 . x_{i2} + {x_2}^2 . {x_{i2}}^2 + 2 x_1 . x_{i1} + 2 x_2 . x_{i2}$$

(12)

$$\varphi(x) = [1, {x_1}^2, \sqrt{2} x_1 x_2, {x_2}^2, \sqrt{2} x_1, \sqrt{2} x_2]^T$$

(13)

Çekirdek fonksiyonları karşılaştırıldığında polinom ve radyal tabanlı kernellerin daha sade ve anlaşılabilir olduğu ifade edilebilir. Matematiksel olarak basit görünse de, polinomun derecesindeki artış algoritmanın karmaşık bir hal almasına neden olmaktadır. Bu da hem işlem süresini önemli ölçüde artırmakta hem de bir noktadan sonra sınıflandırma doğruluğunu düşürmektedir. Buna karşın radyal tabanlı fonksiyonun kernel boyutu (γ) olarak ifade edilen parametresindeki değişimlerin sınıflandırma performansına etkisinin daha az olduğu görülmüştür [5]. Farklı özelliklere sahYP problemlerin çözümünde farklı çekirdek fonksiyonlarının üstünlükleri görülmüştür.

4. DENEY SONUÇLARI

Bu çalışma [6]'deki lingspam_public isimli mail veri seti kullanılarak gerçekleştirilmiştir. Çalışılan eğitim setinin sonuçlar üzerinde belirleyici etkisi olduğu tespit edilmiştir. Normal posta(np) sayısının YP sayısından fazla tutulması kullanıcıya gelen maillin tespitinde np olma ihtimalini arttırmaktadır. Daha önce yapılmış olan çalışmaların çoğunda np sayısı fazla tutmuştur. Bu durumun en uygun sonuçları verdiği gözlemlenmiştir.

DVM ile sınıflandırma işlemine başlamadan önce öznitelik çıkartma ve seçme işlemleri için bazı ön işlemler yapılmıştır; 1. Adımda 3000 np, 1000 YP mesajı kelimelerine parçalanmış ve tüm mesajlar içerisinde en az 5 defa tekrarlayan kelimeler veya özel karakterler seçilmiştir. 2. adımda kelimelerin ve özle karakterlerin olasılık değerleri bulunmuştur. 3. Adımda 96 adet mesaj 2. adımdaki filtreden geçirilerek 4. Adıma ulaşmıştır. Artık her mesajın (0.011 -0.999) arasında bir değeri vardır ve mesajların sınıfları belirlidir. Bu mesajlar eğitim sınıfını temsil etmektedir. Sonuçta, 5.adımda 96 adet test verisi 4.adımdaki filtreden geçirilmiş ve sonuçlar elde edilmiştir.

Önceki çalışmaların hemen hepsinde mesaj içerisindeki özel karakterler sınıflandırma işlemelerine dahil edilmemiştir. Bu çalışmada ise "! # $" karakterleri öznitelik kümesine dahil edilerek ikinci bir sınıflandırma işlemi yapılmıştır. Tablo 2 de görüldüğü gibi '%*?' karakterleri orta derece bir olasılık sonucu verdiği için sınıflandırma işlemine dahil edilmemiştir. Ayrıca, '( )' karakterleri de test işlemine tabi tutulmuş fakat test sonucunda kayda değer bir başarı göstermemiştir ve bu nedenle öznitelik kümesine dahil edilmemiştir. DVM ile sınıflandırma işlemi yapılmış ve her mesajın YP veya np etiketi bulunmuştur. Daha sonra özel karakterler dahil edilerek DVM ile tekrar sınıflandırma yapılmıştır.

Tablo 1: Olasılık Dağılımı ve Başarı yüzdesi

	Öznitelik Kümesi	Test1	Test2	Test3	Test4
Doğru tanıma sayısı	1.Sınıflandırma	89	93	91	89
	2.Sınıflandırma	91	95	95	90
Başarı oranı	1.Sınıflandırma	0.927	0.968	0.947	0.927
	2.Sınıflandırma	0.947	0.989	0.989	0.937

Tablo 2: Özel Karakterlerin Olasılık Dağılım Tablosu

Karakterler	!	#	$	%	(	)	*	?
Olasılık Değerleri	0.9446	0.8709	0.8894	0.5874	0.1524	0.1519	0.5070	0.4367

Tablo 1' de sınıflandırma sonuçları verilmiştir. Sınıflandırma işlemleri sonunda her iki sınıflandırma sonucu karşılaştırılmış sınıflar farklı ise öznitelik sınıfına '!#$' karakterleri dahil edilerek yapılan sınıflandırma sonuçları doğru olarak kabul edilmiştir. Tablo 1'de de görüldüğü gibi her test soncunda özel karakterler dahil edilerek sınıflandırma yapılmış ve tanıma oranının daha yüksek çıktığı gözlemlenmiştir.

5. SONUÇ

Bu çalışmada, bir e-posta kümesi DVM'nin doğrusal ayrılamama durumu için sınıflandırılmış ve daha sonra önceki çalışmalarda öznitelik kümesinin dışında bırakılan bazı özel karakterler öznitelik kümesine eklenerek tekrar sınıflandırma işlemi yapılmıştır. Fark oluşması durumlarında ikinci sınıflandırma doğru kabul edilmiş ve kayda değer oranda performans artışı olduğu görülmüştür. Sınıflandırma işlemleri DVM'nin doğrusal ayrılamama durumundaki gauss çekirdek fonksiyonu kullanılmıştır ve çekirdek boyutu(γ) 1 olarak alınmıştır.

KAYNAKLAR

[1] C.Altunyaprak, "Bayes Yöntemi Kullanılarak İstenmeyen Elektronik Postaların Filtrelenmesi", YL Tezi, Bilgisayar Mühendisliği Bilimleri, Muğla Üniversitesi, Muğla, 2006.

[2] Ü. Aydoğan, "Destek Vektör Makinalarında Kullanılan Çekirdek Fonksiyonların Sınıflama Performanslarının Karşılaştırılması",YL Tezi, Biyoistatistik Anabilim, Hacettepe Üniversitesi, Dalı, İstanbul, 2010.

[3] N. Cristianini and J. S. Taylor, , *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge: Cambridge University Press, 2000

[4] C. Xiao-li , L. Pei-yu , Z. Zhen-fang and Y. Qiu," A method of spam filtering based on weighted support vector machines", IEEE International Symposium on IT in Medicine & Education ITIME '09. 2009, pp. 947- 950

[5] (2013, Nisan), A Practical Guide to Support Vector

Classification,Available:http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
[6] (2013, Nisan), Ling-Spam data set, Available: http://csmining.org/index.php/ling-spam-datasets.html
[7] M.R. Islam, M.U. Chowdhury, W. Zhou," An Innovative Spam Filtering Model Based on Support Vector Machine", Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on 2005, pp. 348-353.
[8] P.N. Tan, M. Steinbach, V. Kumar, *Introduction to Data Mining, Pearson Education*. Indiana: Addison-Wesley, 2006.
[9] S. Tolun, "Destek Vektör Makineleri: Banka Başarısızlığının Tahmini Üzerine Bir Uygulama", Doktora Tezi, İşletme Bölümü, İstanbul Üniversitesi, İstanbul., 2008.
[10] Vapnik, V. (1995) *The Nature of Statistical Learning Theory*, NewYork: Springer-Verlag, 187.
[11] J. Zhiyang, L. Weiwei, G. Wei, X. Youming, "Research on Web Spam Detection Based on Support Vector Machine", Communication Systems and Network Technologies (CSNT) 2012, , pp.517- 520.
[12] Z. Wang, X. Sun, X. Li, D. Zhang,, "An Efficient SVM-Based Spam Filtering Algorithm", Machine Learning and Cybernetics, International Conference on 2006 , pp. 3682- 3686